





Db	522	enacphlsgrfhgrrgggktheeedvhyeqvirrlskreaivvllaqhpvvfvssgne11	581
Oy	541	LFAFGINAAQNNHNFVLAGRRNVLQQIIEPQAMELAFAASKREVEELFNSODESIFFPGR	600
Dd	582	lfafginaqnunhenflagrernvlqgliepqamelaifaaprkeveesfnsgdsilffpgpr	641
Oy	601	OHQOQSFRSTKQOOPLVSLIDFYGF	625
Db	642	qhngqsprstkqgpplvsildtfvf	666
RESULT 3			
ID	W62829	standard; Protein; 666 AA.	
XX	W62829;		
AC	W62829;		
XX			
DT	27-OCT-1998	(first entry)	
XX			
DE	Macadamia integrifolia antimicrobial protein.		
XX			
KW	antimicrobial protein; infestation; control.		
XX			
OS	Macadamia integrifolia.		
XX			
FH	Key	Location/Qualifiers	
CC	Peptide	1..28	
FT		/note= "signal peptide"	
FT	Protein	29..666	
FT		/note= "mature protein"	
XX			
PN	WO9827805-A1.		
XX			
PD	02-JUL-1998.		
XX			
PP	22-DEC-1997;	97WO-AU00874.	
XX			
PR	20-DEC-1996;	96AU-0004275.	
XX			
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.		
XX			
PI	Bower NI, Goulter KC, Green JL, Mannens JM, Marcus JP;		
XX			
DR	WPI: 1998-377279/32.		
DR	N-PSDB: V42311.		
XX			
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -		
XX	useful for controlling microbial infestations of plants or mammals		
XX			
PS	Claim 1; Page 39-41; 96pp; English.		
XX			
CC	The sequence is that of an antimicrobial protein which can		
CC	be used to control microbial infestations in plants and mammalian		
CC	animals.		
XX			
SQ	Sequence 666 AA:		
Query Match 96.7%; Score 3215; DB 19; Length 666;			
Best Local Similarity 96.6%; Pred. No. 2,1e-286;			
Matches 604; Conservative 7; Mismatches 14; Indels 0; Gaps 0.			
OY	1	OCMOLETSGOMRRCYSODCKRFPEEDIDMSKYDNORDPQTECOOCORCHQOESDPOOQY	60
Dd	42	qcmqlctsggmrtcvsgcdkrtfeedidwskydhgdppjdcqqcqrtrcgsgspiqgy	101
OY	61	CORCKEICEEEEEEYNRODPQOQYECCORCQRRETEPRHMOICQORCERYEKERKO	120
Dd	102	cqrckceiceeeeynrqpdygegcqercqrnheteprhmqtcqrcreryeketrky	161
OY	121	OKRYVEDQOREDEBKFEENRKESGDNRDRDQQRTEECORHCFOEQERRLYOCORCOEOOR	180
Dd	162	qkriveeqredeekyeetmkednkrdpqrveyedcirrecedepdqyqcqtrcrerqqr	221

QY	181	QHRGGDLMPNRGGSSGRTEEGEEKQSNPIFYEDBRSLSTRFRFEESHIVLENFGRSK	240
Db	222	ghrvgddlhhpprgsgsrteeegeeksqndpyfiderlsstrfteeghlsvienfygrsk	281
QY	241	LRLALKNNRLVLEENPNPAFLPTPHLDADALLVIGGKGLKMHHPNRRESYNLEGGDVI	300
Db	282	llrlaknrvlvileenpnafvlpchldadalllvtrggrgalkmhndresynleegdv	341
QY	301	RIPAGTFEYLNRDNNERLHIAKFLQTLSPNGOYKEEFPFPAAGGNDEPYLSTFSKEILEAA	360
Db	342	ripagtfylhordnerlhiaqlfqlstprgqkeffpaggqpepylstfakellaa	401
QY	361	LNTQTERLRGLVGOOREVITIRASQEOIRLRTDDSSRRMHIRRGCESSRGYNLFNKR	420
Db	402	lntqterlrvlgvqgrevvltasseqvirelfrddsserrwhlrrgessrgpynlfmkr	461
QY	421	PLYSKRYGOAEVKEPEEDYRQLODDDVSVFINIRIGSGMMGPFENRSTRKYVVVAASGEADV	480
Db	462	plyskrygqvgyevkpeedyrtqldmdvsvflantqsgmmgpfenlrstkyvvvaasgeadv	521
QY	481	EMACPHLSGRHGGKGRGGRHBEEDVEVHYEQVRLSKREALIVYLACHPPVVFVSSGNENLL	540
Db	522	emacphlsgrhnggrgrgkrheeedvhyeqvkarlskrealivpyghpvrvfvasgnea	581
QY	541	LFAEFINQNNHNFENLAGRENVLQOIEPQAMELAFASKREVEBELFNSODESITFFPCPR	600
Db	582	lfaifnqnnhnenflagrenvltqiepqamelafaprkeveelfnsqdesitffp	641
QY	601	QHOOQSPRSTKOQOPLYVSLDPFGE	625
Db	642	qhooqssrstkqgqplvslldfvgf	666
RESULT	4		
R20181			
XX	ID	R20181 standard; Protein; 566 AA.	
AC		R20181;	
XX	DT	16-APR-1992 (first entry)	
XX	DE	Sequence encoded by 67 kD T. cacao protein cDNA.	
XX	KW	Cocoa; flavour; vicilin; seed storage protein.	
XX	OS	Theobroma cacao.	
XX	PN	W09119801-A.	
XX	PD	26-DEC-1991.	
XX	PE	07-JUN-1991; 91WO-GB00914.	
XX	PR	11-JUN-1990; 90GB-0013016.	
XX	PA	(MRSC ) MARS UK LTD.	
XX	PI	Spencer ME, Hodge R, Deakin EA, Ashton S;	
XX	DR	WPI: 1992-024418/03.	
XX	DR	N-PSDB; Q20377.	
PT		Recombinant cocoa proteins - are responsible for flavour in cocoa	
PT		beans and produced in large quantities using yeast and bacterial	
XX		expression vectors	
PS		Claim 4; Fig 2; 59pp; English.	
CC		The inventors claim a 67 kD and 31 kD T. cacao protein, and	
CC		fragments, and encoding DNAs. The 47 kD and 31 kD proteins are	
CC		derived from the 67 kD precursor. T. cacao protein cDNA was	
CC		detected in a cDNA library prepared from immature cocoa beans RNA	





```
Db 320 ghqqlveadarsfhndiaehdvsvsfantiagsmsaplifntrsfklayvpngkyacivcp 379
Oy 486 HLSGRHG-----GRGGKRIEEEEVH-----YEOVRARLSKREAIIVLAGHPVFE 531
Db 380 hqsggggeereerdgyrseeeeeeeseegaeaggyhtrirarisgtafvvpaqhpfa 439
Oy 532 VSSGENILLFAFGINAONNHENFLAGREBNVLOQIEPQAMEIAFASKEVEELFNSOD 591
Db 440 vasrdanlqivcfvhadrnekvflagad-nvlpkldrvakalsfaskaeedevlgsrr 498
Oy 592 ESIFFPGRQ---HOQSPRSTKQOQ 614
Db 499 ekglfpgeesgghneeregeereee 524

RESULT 8
W62837 ID W62837 standard; Protein: 637 AA.
XX AC W62837;
XX DT 27-OCT-1998 (first entry)
XX DE Hordeum vulgare antimicrobial protein.
XX KM antimicrobial protein; infestation; control.
XX OS Hordeum vulgare.
XX PN WO9827805-A1.
XX PD 02-JUL-1998.
XX PF 22-DEC-1997; 97WO-AU00874.
XX PR 20-DEC-1996; 96AU-0004275.
XX PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX DR WPI: 1998-377279/32.
XX PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX PS useful for controlling microbial infestations of plants of mammals
XX PS Claim 1; Page 60-62; 96pp; English.
XX CC The sequence is that of an antimicrobial protein which can
XX CC be used to control microbial infestations in plants and mammalian
XX CC animals.
XX SQ Sequence 637 AA;

Query Match 26.0%; Score 864.5; DB 19; Length 637;
Best Local Similarity 35.0%; Pred. No. 1e-70;
Matches 221; Conservative 100; Mismatches 219; Indels 91; Gaps 19;

Oy 29 SKYDNOEDPQ--TECOQCCRCROQESDPRQOQYOCRKKECEEEENRORDPOQOYE 86
Db 27 ashdaddtrgrghslqgcvgcrger--pr---ysharvcgc-----rddqgqh- 71
Oy 87 OCOKRCORRETPRNMOLCOQRCERRYEKKRKQKRYEEQOREDEDEKYEENKEGDNKR 146
Db 72 -----grhegeeggrgrgwhggegereehgrgrhgegeehgrgrgh 119
Oy 147 DPOQREYEDCRRHCEQOEBRLIOYOCORCQEQORQHRG---GDLMPQSGSGRYEBG 202
Db 120 gegereregrghgrhgege-----reegergrgrhgegeereeggrgrgrgeg 170
Oy 203 E--EKQSDN--PYFDENSLSTRPTEBCHISVLENFYRSKLLRLAKNRYLLEANPN 258
```

```
Db 171 erdeegdsrrpyvlgtrfrtrllgsdnhfvralrpfddvstrllrgldryvaimevnp 230
Oy 259 AFVLPRLHLDADAILLYIGRGALKMIRHNDRESYNLECGDVRIRIPAGTFYILNRDNMR 318
Db 231 atfvpgfdadvgvyvaggvylvlengekrsylvkegdvavapagslmhlanldgrtk 290
Oy 319 LHIAKFLQITISTPGQYKEFPFAGGQNPFFYLSFESKEILEALMOTERLRGVLGQ--- 375
Db 291 lviakllhltisvpgkt-qfl-----svkpllaslskrvlraafktsderlerlfnrgq 344
Oy 376 ---RECVIIRASQEQIRLETRDDSE---SRKWHIRRGSSSGPVLVLFKRPRLYSKTKYO 429
Db 345 ektrsvsivrtaaseqlrelrtaaeaggyghrwpilpfrgdsrtdtllleqrpklanhgr 404
Oy 430 AYEKREDFYROLQMDVSVFIANITGSMKGFPEFNTSRKVVVVASGEADVEMACPHL-- 487
Db 405 lyeadarsfhlanqdvtravaniipgsmtapylnqsfklavvleagevgqlvcpnigr 464
Oy 488 ---SGRHGGRG-----GGRKHEEEEVH-EQVRARLSKREAIIVLAG 526
Db 465 eseserehgykgyrrreeeddqgrqrrrgseseseeegqryelvtvratvrsafvppg 524
Oy 527 HPVVFVSS--GNENLLFAFGINAONNHENFLAGREBNVLOQIEPQAMEIAFASKEVE 584
Db 525 hpvvelssqgsnqlqvclfeinaerwvlagr-nvvlqklgspegelftgrparevq 583
Oy 585 ELFNSODESI-FPPGPRHQOQSPRSTKQOQ 614
Db 584 evfragdqdegfvagp---eqgsregegege 611

RESULT 9
Y15244 ID Y15244 standard; Protein: 626 AA.
XX AC Y15244;
XX DT 09-NOV-1999 (first entry)
XX DE Peanut allergen, Ara h 1, amino acid sequence.
XX KM allergy; immune response; transgenic; allergen; epitope;
XX OS Immunoglobulin E; Ig E; binding site; peanut.
XX OS Arachis hypogaea.
XX PN WO9938978-A1.
XX PD 05-AUG-1999.
XX PF 29-JAN-1999; 99WO-US02031.
XX PR 27-AUG-1998; 98US-0141220.
XX PR 31-JAN-1998; 98US-0073283.
XX PR 13-FEB-1998; 98US-0074590.
XX PR 13-FEB-1998; 98US-0074624.
XX PR 13-FEB-1998; 98US-0074633.
XX PA (SOSI/) SOSIN H.
XX PA (UYAR-) UNIV ARKANSAS.
XX PA (UNWV) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
XX PI Bannon GA, Burks AW, Sampson HA, Sosin H;
XX DR WPI: 1999-479189/40.
XX DR N-PSDB: 206382.
XX PT Modified allergen with reduced IgE binding, useful for treating e.g.
XX PS allergies
XX PS Disclosure: Page 35-37; 46pp; English.
```







XX W09827805-A1.  
 XX 02-JUL-1998.  
 XX 22-DEC-1997; 97WO-AU00874.  
 XX 20-DEC-1996; 96AU-0004275.  
 XX (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 XX WPI; 1998-377279/32.  
 XX Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 XX useful for controlling microbial infestations of plants or mammals  
 XX Claim 1; Page 55-57; 96pp; English.  
 XX The sequence is that of an antimicrobial protein which can  
 XX be used to control microbial infestations in plants and mammalian  
 XX animals.  
 XX  
 SQ Sequence 614 AA;

Query Match 25.9%; Score 861; DB 19; Length 614;  
 Best Local Similarity 34.5%; Pred. No. 2e-70;  
 Matches 220; Conservative 114; Mismatches 205; Indels 98; Gaps 23;

QY 29 SKYDNEDEPQTEC-QQCQRRCROESDPRQOQCQRCKEICEEEDYRNRPDQOQED 87  
 D 27 spyrktemp---cagrcqlqsc-qgepddlkqkacesrckl-----ey-----dprecvtd- 72  
 QY 88 CQKRCQRRETERPHNMICQQRCKRREKREKQKQKRYEQRDEKRYERMEKED--NK 145  
 D 73 -----tgalnqhrrp-----getrtgrrpdy-----dddrtrprtreagrrwqr 111  
 QY 146 RDPQREYEDCRHRCQDEPRLOYOCORRCQEQ-----RQHGSGGLMNPQRCGSRYE 201  
 D 112 aeprereere-----dwqrpredw-----trpsbqprkllpregrege-----qewqrrpsev 159  
 QY 202 GEEKQSDNPYFEDSLSTRFTEGHSIVLENFYGRSKLLRALKNYRLVLEANPAEV 261  
 D 160 reetstnpfyfrrfstrfyngnrrlrvlqfdrskqfqnlnhrivqleaprrtlv 219  
 QY 262 LPHILADAILLVIGRGALKMTHRNRESYNIECGDVIRIPAGTFFYLINRDNERRLI 321  
 D 220 lphhadadhlvlyqgqatvlyvngnrrksfnldeghalrpsgfysyllnrhdnqlrv 279  
 QY 322 AKPLQTIPTPGQKKEFPFAGGONPEPYLSTFSKEILEALNQTRELKRGV----- 372  
 D 280 aksmpvnrpqrqfedfrrpsstrdgsylyqfscnlleaefneirrvlleenagqev 339  
 QY 373 ---GQGR-----EGVIRASQEQIRBLTRDSESRMHIRRGESS--RGRVYLFNK 419  
 D 340 eergqrrrtstresdnegvlyvkshvqgeltkhakevs-----kkgseeditnplnldrg 395  
 QY 420 RPLYSKKYGQAVEVRKEDR-QLQDMDVSVFIANTIGSGMGRFRTTRTKYVVVAVASGA 478  
 D 396 epqlsnmfrllvkrpdkknprqldmmltcvelkegalmlrphnlskamvlyvvnkgtg 455  
 QY 479 DVEMACPHLSGRHGGGKGRHEEEHVE-----QVR---ARLSRRLVILIAGHVVVF 531  
 D 456 nlelvavtrlegqrrgrregeweeedeegsnrevrrtyarlkegdvflmpaaprvai 515  
 QY 532 VSSGNENLLFAFGINAONNHENFLAGRRNVLQQLPEQAMELAFASRKEVEELFNSOD 591  
 D 516 nassehlhl--gfginaemhrrflagdkdvldqtekdaklafrgsgeyekllkngr 573  
 QY 592 ESTFFPG-PROHOOOSPRSTKQO-----PLVSTL 620

DB 574 eshlvsarpgsqspsspekedeengegkyp11s11 610

RESULT 13

W62838 standard; Protein: 605 AA.

W62838;

27-OCT-1998 (first entry)

Glycine max antimicrobial protein.

antimicrobial protein; infestation; control.

Glycine max.

W09827805-A1.

02-JUL-1998.

22-DEC-1997; 97WO-AU00874.

20-DEC-1996; 96AU-0004275.

(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

WPI; 1998-377279/32.

Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 useful for controlling microbial infestations of plants or mammals

Claim 1; Page 63-65; 96pp; English.

The sequence is that of an antimicrobial protein which can  
 be used to control microbial infestations in plants and mammalian  
 animals.

Sequence 605 AA;

Query Match 25.8%; Score 859.5; DB 19; Length 605;  
 Best Local Similarity 32.1%; Pred. No. 2.7e-70;  
 Matches 203; Conservative 133; Mismatches 197; Indels 99; Gaps 18;

QY 31 YQNEDEPQTECQOCQRRCROESDPRQOQCQRCKEICEEEDYRNRPDQOQED 77  
 D 27 ywekenph--nkclqsc-nserdsyngacharcnllkvekecegeglprprprpqh 83  
 QY 78 QRDPOQOYQOCQKRCR-----RETERPHNMICQQRCKRREKREKQKQKRYEQRDE 130  
 D 84 erepprqgkededqrrlprfpqrqeeheqevrkrkqekgseeded 143  
 QY 131 DEKYEERMEKEDNKRDPQREYEDCRHRCQDEPRLOYOCORRCQEQ-----RQHGSGGLMNPQRCGSRYE 190  
 D 144 edeeqder--qfrrfprpqkq-----erreeeded-----eeqres----- 179  
 QY 191 PORGSGRYEGBEERQ-----SDNPYFEDSLSTRFTEGHSIVLENFYGRSKLLRAL 245  
 D 180 -----eeesedslrrhknknpdlfsgnfcelfknqyrrlvqlqfngrrspqln 230  
 QY 246 KMYRLVLEANPAEVLPRLHLDADAILLVIGRGALKMTHRNRESYNIECGDVIRIPAG 305  
 D 231 rylrrlfnfksknclllphnadadyllvlnqta1stsvnndrdsyrtqsgdalrvpsg 290  
 QY 306 TTFYLINRDNERRLIHAKFLQTIPTPGQKKEFPFAGGONPEPYLSTFSKEILEALNQT 365  
 D 291 tlyyvvnpdmnenlllrltalrvpnkgrfresffisstgaqsylyqfscnlleaaydtkf 350  
 QY 366 ERLKRVL-----GQGR---EGVIRASQEQIRBLTRDSESRMHIRRGESSRGPY 414

Db	351	eenkvlfisreagqgqgqrlqesviveiskeqiralsrakssrktl-----sdekp	406
Oy	415	NLFNKRPLRYSNRKOAYEVRKEPDYROLQMDVSVYFANITQSGMMGPFNFRSTKVVYA	474
Db	407	nlrsrdpysnklgkffetpcknqjrrldifisvdmegalllphfnskaiavlvin	466
Oy	475	SGENDVEMACHLSGRIGRCRGCKRHHEEVEVHEQ-----VRALSKREALVYLAGHP	528
Db	467	egdanlelv-----glkeqgqgqgeqglvyrkyraelsedqlfvdpayp	513
Oy	529	VVFSSGNEENLFLFAFGINQNNENENFLRERENVLOQIEPQAMELAFASRKEVELFN	588
Db	514	vvv--natsnlnffaiginaemgnrlfiasqdvvisqpsqvetafpssaqaveklk	571
Oy	589	SDQSEIFPPGPRQHQQSPNSTKQOQPLVSTL	620
Db	572	ngresyfvdaqpkkkkegnkgrk--gplssll	601
RESULT 14			
ID	W22150	standard; Protein; 626 AA.	
XX	AC	W22150;	
XX	XX	29-DEC-1997 (first entry)	
XX	DE	Peanut allergen Ara hi.	
XX	KM	Peanut; seed storage protein; allergen; allergy; hypersensitivity;	
XX	KM	vaccine; anaphylactic shock; immunotherapy; therapy;	
XX	KM	monoclonal antibody; ELISA; analysis; Ara hi.	
OS	XX	Arachis hypogaea strain Florunner.	
XX	XX		
FH	FT	Key	Location/Qualifiers
FT	FT	Peptide	1..22
FT	FT	Protein	/label= Sig_peptide
FT	FT	Modified-site	23..626
FT	FT		/label= Mat_protein
FT	FT		521..523
FT	FT		/note= "N-glycosylation site"
XX	XX	W09724139-A1.	
XX	PD	10-JUL-1997.	
XX	PE	23-SEP-1996; 96NO-US15222.	
XX	XX	04-MAR-1996; 96US-0610424.	
PR	PR	29-DEC-1995; 95US-0009455.	
PA	XX	(UYAR-) UNIV ARKANSAS.	
PI	XX	Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;	
DR	XX	WPI: 1997-363453/33.	
DR	XX	N-PSDB; T76613.	
XX	XX		
PT	XX	Peanut allergens Ara hi and Ara hII - used for vaccination and in	
XX	XX	two-site monoclonal antibody based ELISA	
XX	XX	Claim 31; Page 172; 354pp; English.	
CC	CC	This polypeptide comprises major peanut allergen Ara hi (W22149).	
CC	CC	Its sequence was deduced from cDNA clone P4D (T76613), isolated	
CC	CC	from peanut seed cDNA using a primer (see T76616) based on an	
CC	CC	isolated Ara hi peptide (see W24206). The sequence shows	
CC	CC	significant homology with the vicilin family of seed storage	
CC	CC	proteins of other legumes. The allergen is recognised by serum	
CC	CC	IgE from a large proportion of individuals with peanut	
CC	CC	hypersensitivity. Ara hi and Ara hII (see W24164) can be used to	
CC	CC	raise monoclonal antibodies which are used in a specific two-site	

CC MAb ELISA for the detection of Ara hi or Ara hII (claimed). IgF-  
CC binding Ara hi antigen epitopes (see W24165-87) may be used in  
CC vaccines to protect against allergic reactions to peanut allergens  
CC e.g. anaphylactic shock.

Sequence 626 AA;

Query Match	25.8%;	Score 857.5;	DB 18;	Length 626;
Best Local Similarity	34.3%;	Pred. NO. 4.4e-70;		
Matches 209;	Conservative 113;	Mismatches 209;	Indels 79;	Gaps 18;

```

QY 52 ESDPPOOY---CQRCKEICEEEEEENRKORDPOOYOCQRCRCORETBER-----H 101
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 26 ksspykktlencpaarctqscqgqdddkkq-----acserckleydpvcydrpgh 78
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 102 MOICQNCERRERKEKRRKOOKRYEBEOOREDEKYEERNKQED--NKRPDOOREYDCKRH 159
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
79 lgtlntqisppa-ertrrgqagdy-----dddrtrrteegrgwpaqreteree--- 127
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

160 CEQEPRLQYCCQRKCEEO --- RQHGRCGDLMPQRCGSGRYEEGEEKOSDNPYYFDE 215

```

Db      .128 -dwqrpdw---rrslhqpqrkirepqrqde---qewatpqshvreetsrnmqfyfys 179
      :::|| : || |::| : | : ||::||

```

216 RSLSTREFTTEEGHTSVLENFYGRSKLTBAIKNYRIYLTPANPNAFVLPTHIDADATILLYT 275

```

180 rrfst rrvnanaar i rv] arfdarsra f an] anbr i va i eak n t l v] o k h a d a d n i l v i a 239
      | | | : : | | | : | | : : | | : | | | | | | | | : :
Db

```

00 376 ССБСАКМТННДНБЕСУНІ ЕССДВІРІДАСТМЕУІ ТНБДННБІІАКЕІОТІСТБСОВК 335

[illegible]

```

00: 336 EEEEDACCONDEBYI EMEEXETI EAFI NMECEBI BCUI -----COOB----- 376

```

[illegible][illegible][illegible][illegible][illegible][illegible]

```

430 KGGGGGGGKKE-EELVMI EQVR--AKLSKKEAI VVLAQH FV FVSSGNNELLLFAFG 343
: || : | ||| :|| ||| : : | ||| :| :|| |

```

DB 4/5 qqrgreeeeeeeegsnrevrlycalrkgo.v.lmipaalpvalnaselml--grg 532

QY 546 1NAQNNHENE LAGREKRN LQQIEPQAMELAF AASRKEVEEELNSQDESIFPGPRQHQQ 605

Db 533 lnaennhrlllagdkdnvidqlekqakdlatfpgsgeqvellknqkeshfvsarpgsq 592

QY 606 SPRSTKQQQP 615

Db 593 spsspekesp 602

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

Y40999

[illegible]

0  
h  
t  
t  
p  
:  
/  
w  
w  
w  
.  
e  
l  
c  
o  
n  
g  
r  
e  
s  
s  
.  
g  
o  
v

[illegible]

XX  
be soybean beta conglyutinin protein sequence.

allergic reaction; soybean; beta-conglycinin.

AA  
OS      Glycine max.

AA  
PN W09945961-A1.

```
XX 16-SEP-1999.
PD
XX
XX 12-MAR-1999: 99WO-US05494.
PF
XX
PR 12-MAR-1998: 98US-0077763.
PR 11-MAR-1999: 99US-0077763.
XX
XX (UYAR-) UNIV ARKANSAS.
XX
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS,
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA.
XX
XX WPI; 1999-551218/46.
XX
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction -
XX
XX PS Disclosure; Fig 33A-B; 193pp; English.
XX
XX CC The invention provides a tertiary structure for the peanut allergen
CC Ara h 1. The Ara h 1 allergen is found to contain 23 linear IGE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IGE. The invention may also be used to
CC ensure that the allergen is not introduced into genetically modified
CC food. The present sequence represents a soybean beta-conglycinin protein
CC sequence.
XX
SQ Sequence 605 AA:

Query Match 25.3%; Score 842.5; DB 20; Length 605;
Best Local Similarity 31.8%; Pred No. 9,9e-69;
Matches 203; Conservative 131; Mismatches 195; Indels 109; Gaps 19;

OY 28 WSKYDNOEDPQTECCQCORRCQESDPROQOYCORCKEICEEEFYNNRORDPQOQYEQ 87
   I : : : : : I : I : I : I : I : I : I : I : I :
DB 28 weknkphn-----kclqsc-nserdsyrngcharcnllkveckeckgeiprp---- 76
   I : : : : : I : I : I : I : I : I : I : I : I :
OY 88 COKRCQRRRETERPHNMQICQORCERRYEKEK-----KQCKRYEE--QOREDEE---KY 135
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 77 -rprpqhperp-----qqpgekeedeqpripfpqpqrgeehqreegeewprke 129
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 136 EERMKEGDNKRBDPQOREYEDCRHCEOE-----PRLOYOCORR-----COEQORQHCR 184
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 130 ekrgyegseedeede-----eqqerqfipfripqhkeerneedeedeegqres-- 179
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 185 GGDLMNPQRGSGRYEGEGEKQ-----SDNPYVFDERSLSTRFRTECHISVLENFYGRS 239
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 180 -----eesedseirrhknknpflfgsnrfetlfnkngygrlvyldqrfngts 224
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 240 KLLRAKKNRYLVLEENPNAPFVLPHTLDADAILLVIGRGALKMHRDNRESYNLECGDV 299
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 225 pqqlnrlrdyrllefnskpntlllphnadadyilvlnqtaislvnnddrdsyrlqsgda 284
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 300 IRIPIAGTFYVLINRDNNERLHTAKFLTQISTPQOYKEFPAGOGNEPPLSTFSKELTEA 359
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 285 lrvpsqtltyvvnphnennrlitlaipvnpkprfessflsttraqsylygqfstrniles 344
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 360 ALNTOERLRGLV-----GQOR--EGVIRASQEQIRELTRDSESRRWHLRNGE 408
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 345 aydtkfeelinkvlfsvreegqgqegrlgesvavelskeqiralskrakessrktl----s 400
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 409 SSGRPYLNLFNKNRKYLNKRGQAYEVKPEQYRQLQDMDVSVFANTITQSGMAGPFETRAST 468
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 401 sedkprlrlsrpilynsklyqkfelfekpnqplrdldiflislvmegalllphfnska 460
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 469 KVVVVASGEADVEMACPHLSGRHGRGSGKRHEEVEHYEQ-----VRARISKREALIV 522
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 461 vllvlnegandanelv-----glkegqegqgeeplevrkyraelsedqltv 507
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
```

```
OY 523 VLAGHPVVFVSSGNENILLFAGFINAONNIENFLAGRERNVLOQIEPQAMELAFASRKE 582
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 508 ipagypvvv--natnlnlfaiginaennqgrflagsqdnvlsqipsyqetalfpsaga 565
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
OY 583 VERLEFNSODESTIFFPGPRQHQOOSPRSTKQOQPLVSL 620
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 566 VKKllkngrksyfvdaqpkpkkeegnkgRK--qpIssl 601
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
```

Search completed: March 1, 2001, 15:47:15  
Job time: 240 sec

